Production Workflow Tracking and QC Analysis at the Joint Genome Institute

Heather Kimball, Stephan Trong, Art Kobayashi, Sam Pitluck, Yunian Lou, Matt Nolan
U.S. DOE Joint Genome Institute, Walnut Creek, CA 94598

The Joint Genome Institute Production Genomics Facility has produced over 2.75 billion bases of draft paired-end sequencing since January 1, 2001. Our sequencing methodologies incorporate two types of DNA template generation: inoculation/SPRI purification and Rolling Circle Amplification. In order to manage the flow of samples through these processes, a robust database tracking system was developed using ORACLE. The key elements that are tracked within the workflow system include:

- Instruments
- Operators
- Protocols
- Reagents
- Dates and times
- Quality scores and contamination information

Data input and reporting for the workflow have been produced using a combination of commercial database development software and in-house programs. These include ORACLE's WebDB and Perl CGI programming. By leveraging the rapid report and form development cycle using WebDB and augmenting this with the flexibility of in-house programming, we have efficiently deployed a critical laboratory information management system for our data tracking. This work was performed under the auspices of the U.S. Department of Energy, Office of Biological and Environmental Research, by the University of California, under Contracts No. W-7405-Eng-48, No. DE-AC03-76SF00098, and No. W-7405-ENG-36.